

EXPRESS MAIL NO. EL 669 268 402 US

DATE OF DEPOSIT June 26, 2001

D. J
11-16-01
HG
Priority Papers
U.S. PTO
09/09/01
06/26/01

Our Case No. **9333/274**
Client Reference No. **IW US 01002**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICES

In re Application of:

Kiuchi et al.

Examiner: Pending

Serial No.: New Application

Group Art Unit No.: Pending

Filing Date: Herewith

For: Voice Feature Extraction Device

SUBMISSION OF CERTIFIED COPY OF PRIORITY DOCUMENTS

Assistant Commissioner for Patents
Washington, D.C. 20231

Dear Sir:

Transmitted herewith is a certified copy of priority document Japanese Patent Application No. 2000-212067, filed July 12, 2000, for the above-named U.S. application.

Respectfully submitted,

James P. Naughton
James P. Naughton
Registration No. 30,665
Attorney for Applicant

BRINKS HOFER GILSON & LIONE
P.O. BOX 10395
CHICAGO, ILLINOIS 60610
(312) 321-4200

日本国特許庁
JAPAN PATENT OFFICE

JPO
09/891876
06/26/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application:

2000年 7月12日

出願番号

Application Number:

特願2000-212067

出願人

Applicant(s):

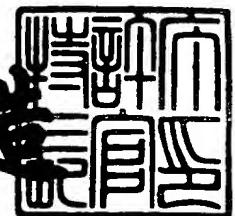
アルパイン株式会社

CERTIFIED COPY OF
DOCUMENT

2001年 5月30日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2001-3047671

【書類名】 特許願
【整理番号】 IWP99290
【あて先】 特許庁長官 殿
【国際特許分類】 G01L 3/00
【発明者】
【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式
会社内
【氏名】 木内 真吾
【発明者】
【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式
会社内
【氏名】 浅野 利明
【発明者】
【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式
会社内
【氏名】 斎藤 望
【特許出願人】
【識別番号】 000101732
【氏名又は名称】 アルパイン株式会社
【代理人】
【識別番号】 100111947
【弁理士】
【氏名又は名称】 木村 良雄
【手数料の表示】
【予納台帳番号】 060750
【納付金額】 21,000円
【提出物件の目録】
【物件名】 明細書 1
【物件名】 図面 1

特2000-212067

【物件名】 要約書 1
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声特徴量抽出装置

【特許請求の範囲】

【請求項1】 予め使用するノイズリダクションシステムのノイズリダクションシステム係数を演算するノイズリダクションシステム係数演算手段と、

被処理入力音声のパワースペクトルベクトルを演算する入力音声パワースペクトル演算手段と、

前記入力音声パワースペクトル演算手段で得られたパワースペクトルベクトルを、前記ノイズリダクションシステム係数演算部で得られた係数に設定したノイズリダクションシステムにより演算処理することを特徴とする音声特徴量抽出装置。

【請求項2】 前記ノイズリダクションシステム係数演算部は、使用するノイズリダクションシステムのフィルタ係数を確定するフィルタ係数算出手段と、前記フィルタ係数算出手段で得られたフィルタ係数をパワースペクトルベクトルに変換するパワー算出手段とからなることを特徴とする請求項1記載の音声特徴量抽出装置。

【請求項3】 前記フィルタ係数算出手段は、入力音声と模擬音声の各信号を加算した信号に対して適応制御を行い、タップ係数を得ることによりフィルタ係数を算出するものであることを特徴とする請求項2記載の音声特徴量抽出装置。

【請求項4】 前記模擬音声信号に対して所定のゲイン調整処理を行うことを特徴とする請求項3記載の音声特徴量抽出装置。

【請求項5】 前記音声特徴量抽出装置を、車両用ナビゲーション装置の音声認識装置に適用したことを特徴とする請求項1記載の音声特徴量抽出装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声認識、話者認識、ラウドネス補償システム等において、周囲のノイズを減少しつつ、パワースペクトラムを用いて音声の特徴量を抽出する音声

特微量抽出装置に関する。

【0002】

【従来の技術】

利用者が発声した音声を認識する音声認識装置、複数の人が話した中から特定の人の音声を識別する話者認識装置、あるいは周囲のノイズを除去するラウドネス補償システム等においては、対象とする特定の音声を抽出するため、マイクから入力した音声のパワースペクトラムを求め、特定の音声を抽出する処理が行われている。

【0003】

このような特定の音声を抽出するためのシステムを採用する一分野としての音声認識装置における、例えば車両用ナビゲーション装置における目的地設定などを音声によって入力する音声認識装置にこれを用いる際には、車室内にはエンジン音、タイヤの摩擦音、風切り音、更にはオーディオ音等種々のノイズが存在する。そのような中でマイクに入力される利用者の音声は、ノイズによりS/N比が低下して不明瞭な音声となってしまい、このような状態では精度良く音声の特微量をとらえて正確な音声認識を行うことができなくなる。したがって、周囲に大きな雑音が存在する環境下で音声認識装置を使用する際には、利用者の音声を周囲の雑音と分離して処理することが、音声認識率を高める上で重要な要件となっている。このようなことは車両用音声認識装置に限らず、種々の分野で用いられている音声認識装置において問題となるところであり、更に音声認識装置に限らず、上記のような種々の音響技術分野で問題となる。

【0004】

このようなとき、従来の装置においては例えば図3のブロック図に示すように、マイク30から入力した周囲の雑音も含めた音声X(n)を、タップ長がNのFIRフィルタを備えたノイズリダクションシステム(NRシステム)31に入力する。このノイズリダクションシステムでは、サンプリング周波数をf_s(Hz)としたとき、 $1/f_s$ (秒)毎に動作を行う。

【0005】

その後ここで得られた $x(n) * w(n)$ の信号を、ハニング窓やハミング窓

32の窓関数部を通し、サンプリングした1フレーム毎の間のデータ飛躍による高周波成分の発生を防止し、これをL点について処理するFFT演算部33に入力する。このFFT演算部33では、1フレーム毎に高速フーリエ変換を行い、時間軸信号を周波数成分に変換する。ここで得られたスペクトル信号X(n)・W(n)をパワー算出部34に入力し、パワースペクトルの計算を行い、信号[X(w)]₂・[W(w)]₂を得ている。

【0006】

【発明が解決しようとする課題】

上記のような従来の音声特徴量抽出装置においては、このノイズリダクションシステム31でのサンプリング周波数をf_s(Hz)とし、FIRフィルタのタップ長をNとしたとき、1回の処理を行うのにN回の乗算と(N-1)回の和算を行う必要がある。その後、この特徴量を求めるためにFFT演算部33に入力して処理を行うので、このときのFFT演算部33の処理点数をL点とすると、L/2次の特徴量ベクトルを出力するためには、ノイズリダクションの出力がL点必要になる。そのため、結局このノイズリダクションシステムではN×L回の乗算と、(N-1)×L回の和算を行う必要がある。

【0007】

このような処理を行うため、より正確な音声特徴量を抽出しようとして前記NやLの値を大きく設定するとその演算処理量が莫大となり、この装置におけるプロセッサへの負担が大きくなる。そのため、処理速度も遅くならざるを得ず、時には他の処理にも影響を与え、システムの円滑な処理を行うことができなくなる。また、前記NやLの値を小さくとると、当然正確な音声特徴量の抽出は行うことができない。

【0008】

したがって本発明は、音声特徴量の抽出に際して従来と同様の正確な抽出性能を維持しつつ演算処理量を削減し、プロセッサへの負担を軽くし、その処理速度を上昇させることができるようにした音声特徴量抽出装置を提供することを目的とする。

【0009】

【課題を解決するための手段】

本発明は上記課題を解決するため、予め使用するノイズリダクションシステムのノイズリダクションシステム係数を演算するノイズリダクションシステム係数演算手段と、被処理入力音声のパワースペクトルベクトルを演算する入力音声パワースペクトル演算手段と、前記入力音声パワースペクトル演算手段で得られたパワースペクトルベクトルを、前記ノイズリダクションシステム係数演算部で得られた係数に設定したノイズリダクションシステムにより演算処理することを特徴とする音声特徴量抽出装置としたものである。

【0010】

また、前記ノイズリダクションシステム係数演算部は、使用するノイズリダクションシステムのフィルタ係数を確定するフィルタ係数算出手段と、前記フィルタ係数算出手段で得られたフィルタ係数をパワースペクトルベクトルに変換するパワー算出手段とからなることを特徴とする前記音声特徴量抽出装置としたものである。

【0011】

また、前記フィルタ係数算出手段は、入力音声と模擬音声の各信号を加算した信号に対して適応制御を行い、タップ係数を得ることによりフィルタ係数を算出するものであることを特徴とする前記音声特徴量抽出装置としたものである。

【0012】

また、前記模擬音声信号に対して所定のゲイン調整処理を行うことを特徴とする前記音声特徴量抽出装置としたものである。

【0013】

また、前記音声特徴量抽出装置を、車両用ナビゲーション装置の音声認識装置に適用したものである。

【0014】

【発明の実施の形態】

本発明の実施例を図面に沿って説明する。図1は本発明による音声特徴量抽出装置の基本的な機能ブロック図であり、前記従来の装置と同様の機能ブロック部分は同様に示している。同図において、例えば本発明による音声特徴量抽出装置

が車両用ナビゲーション装置に適用される場合には、マイク1から目的地設定等の種々の操作用、指示用の音声や、それに伴って入力される周囲のエンジン音、タイヤの摩擦音、風切り音等の周囲の騒音が、信号 $x(n)$ として窓関数部2に入力される。

【0015】

したがって、本発明における「被処理入力音声」は、この実施例においては利用者の音声と周囲のノイズを含んだ信号 $x(n)$ となる。この被処理入力音声は窓関数部2において、前記従来例と同様にハニング窓やハミング窓の窓関数を通してにより、サンプリングした1フレーム毎の間のデータ飛躍による高周波成分の発生防止等の処理がなされる。

【0016】

これをL点について演算するFFT演算部3に入力する。このFFT演算部3では、1フレーム毎に高速フーリエ変換を行い、時間軸信号を周波数成分に変換する。ここで得られたスペクトル信号 $X(w)$ をパワー算出部4に入力し、 $L/2$ 次のパワースペクトルを算出し出力信号 $[X(w)]^2$ を得る。本発明においては、このようにして得られた入力音声の特徴量である $L/2$ 次のパワースペクトルベクトルに対して、ノイズリダクションシステム5によって従来と同様のノイズリダクション処理を行うこととなる。

【0017】

このノイズリダクションシステム5においては、前記のように処理されて入力されるパワースペクトル信号は $L/2$ 次であるので、このシステムにおいて使用される係数値は $L/2$ 個となり、したがってここで計算される W_m における m の値は0から $L/2$ まで行えば良いこととなる。

【0018】

このノイズリダクションシステムでの処理において、前記のように $L/2$ 次の特徴量ベクトルを出力する場合は、

入力される音声のパワースペクトルを

【数1】

$$\vec{P} = (P_0 \ P_1 \ \dots \ P_{L/2})^T \quad \dots \dots \quad (1)$$

後に説明する予め計算されたノイズリダクションシステムの係数を、

【数2】

$$\vec{w} = (W_0 \ W_1 \ \dots \ W_{L/2})^T \quad \dots \dots \quad (2)$$

としたとき、

出力される特徴量ベクトル

【数3】

$$\vec{Y} = (Y_0 \ Y_1 \ \dots \ Y_{L/2})^T \quad \dots \dots \quad (3)$$

は、

【数4】

$$\vec{Y} = (\vec{P})^T \times \vec{w} \quad \dots \dots \quad (4)$$

となる。

【0019】

このときのノイズリダクションシステムの演算量は、 $L/2$ 回の乗算となるため、前記図3に示す従来の音声特徴量抽出装置における、NタップのFIRフィルタによる時間領域のノイズリダクション処理と比較して、 $1/(4N-2)$ の演算量で済むこととなる。

【0020】

上記のノイズリダクションシステムの特性を決定する係数については、例えば図2に示すようなノイズリダクション係数算出システムによって予め得ておくこ

とができる。ここで用いられるCNR係数を求めるシステムのブロック図においては、従来から用いられている1chの音声強調システムを応用した例を示しており、それにより比較的簡単なシステムにより前記係数算出を行うことができる。

【0021】

即ち、このシステムにおいては、マイク10で周囲の音声を入力する一方、所定の模擬音声信号を発声する模擬音声発声部11からの信号を入力してこれにゲイン調整部12でゲイン調整を行い、これを加算器13でマイク10の信号と加算し、一方、ゲイン調整後の信号を遅延処理部14でm段（L/2段）の遅延を行い信号d(n)を得る。前記加算器13で加算された信号をタップ係数Wの適応フィルタ15で処理し信号y(n)を得る。この信号と前記遅延処理部14からの信号d(n)とを減算器16にかけ、その誤差信号e(n)によりタップ係数Wを調整している。このような制御により得られたタップ係数の値を、この適応フィルタの係数、即ちこのノイズリダクションシステムのフィルタ係数 W_{CNR} として求める。上記処理部分が、本発明におけるフィルタ係数算出手段に相当する。

【0022】

ここで求められる適応フィルタのタップ長は、特徴量を求める際の前記FFTの処理点数と同じに設定し、このようにして得られたこのノイズリダクションシステムの特性を示す適応フィルタ係数 W_{CNR} に対してFFT演算部17でFFT処理する。これをパワー算出部18に入力し、前記ノイズリダクションシステムの係数を得る。上記処理部分が本発明におけるフィルタ係数をパワースペクトルベクトルに変換するパワー算出手段に相当し、前記フィルタ係数算出手段とのパワー算出手段によって、本発明のノイズリダクションシステム係数演算手段が形成されている。

【0023】

なお、図2に示すシステムにおいて、ゲイン調整部12のゲインは、適応フィルタの出力y(n)の歪みとSN比改善量とのトレードオフによって決められる。即ち、ゲインを大きくすると、歪みは小さくなる反面、SN比の改善量は小さ

くなり、逆にゲインを大きくすると、歪みは大きくなるがS/N比の改善量が大きくなる。したがって使用するシステムの特性に合わせて、前記ゲイン調整部12のゲインを調整する。

【0024】

このようにして、予め時間領域でノイズリダクションシステムの特性を求め、それにより予め確定してたフィルタ係数 W_{CNR} をFFT処理し、パワーを算出することによりパワードメインでのノイズリダクション係数ベクトルが求まり、それにより予め前記式(2)に示すようなL/2次のノイズリダクションシステムの係数が求められることとなる。

【0025】

図1に示すノイズリダクションシステム5においては、パワー算出部4で得られた前記式(1)に示すような音声のパワースペクトルベクトルに対して、前記のように求められた式(2)のノイズリダクションシステム係数を用いたノイズリダクションシステム5により、式(4)の演算処理を行うことによって特徴量ベクトルを得ることができる。

【0026】

上記のように、本発明による音声特徴量抽出装置においては、予めノイズリダクションシステムの特性を表すノイズリダクション係数を求めておき、音声処理の最初にFFT処理してパワースペクトルを得て、これに対して前記ノイズリダクションシステムで処理するため、従来の装置のように最初にノイズリダクションシステムで処理し、これをFFT処理するもののように入力音声処理時の位相部分の処理が不要となり、その処理負担は従来の装置の1/(4N-2)の処理量となり、しかも精度は従来のものとほとんど変わることがない高精度の音声特徴量抽出を行うことができるようになる。

【0027】

このような音声特徴量抽出装置は、上記のような車両用ナビゲーション装置の音声認識装置に限らず、各種音声認識装置、更には話者認識装置、ラウドネス補償システム等に広く用いることができる。

【0028】

【発明の効果】

本発明は上記のように構成したので、音声特徴量の抽出に際して従来と同様の正確な抽出性能を維持しつつ演算処理量を削減し、プロセッサへの負担を軽くし処理速度を上げることができる。

【0029】

また、ノイズリダクションシステム係数演算部が、使用するノイズリダクションシステムのフィルタ係数を確定するフィルタ係数算出手段と、前記フィルタ係数算出手段で得られたフィルタ係数をパワースペクトルベクトルに変換するパワー算出手段とからなるものにおいては、予め使用するノイズリダクションシステムのノイズリダクションシステム係数を簡単な手法により容易に、且つ正確に求めることができる。

【0030】

また、フィルタ係数算出手段が、入力音声と模擬音声の各信号を加算した信号に対して適応制御を行い、タップ係数を得ることによりフィルタ係数を算出するようにしたものにおいては、予め使用するノイズリダクションシステムのノイズリダクションシステム係数を、従来から既に1chの音声強調システム等で用いられている方式を用いて容易に、且つ正確に求めることができる。

【0031】

前記模擬音声信号に対して所定のゲイン調整処理を行うものにおいては、そのゲインを大きくすると、歪みは小さくなる反面、S/N比の改善量は小さくなり、逆にゲインを大きくすると、歪みは大きくなるがS/N比の改善量が大きくなるので、使用するシステムの特性に合わせてこれを調整することにより、所望の特性の音声特徴量抽出装置を得ることができる。

【0032】

また、前記音声特徴量抽出装置を車両用ナビゲーション装置の音声認識装置に適用したものにおいては、エンジン音やタイヤの摩擦音、更には風切り音等、極めてノイズの多い音響上悪い環境条件下において、しかも大型の高価な演算処理装置を採用することが困難であり、且つ例えばナビゲーション装置の経路案内処理等にもその演算処理装置を併用しなければならないような環境下において、本

発明を有効に利用することができる。

【図面の簡単な説明】

【図1】

本発明による音声特徴量抽出装置に用いられる、主要処理部分を示す回路構成図である。

【図2】

本発明に用いられるノイズリダクションシステムにおいて、予めそのノイズリダクションシステム係数を求める例を示す回路構成図である。

【図3】

従来の音声特徴量抽出装置に用いられる、主要処理部分を示す回路構成図である。

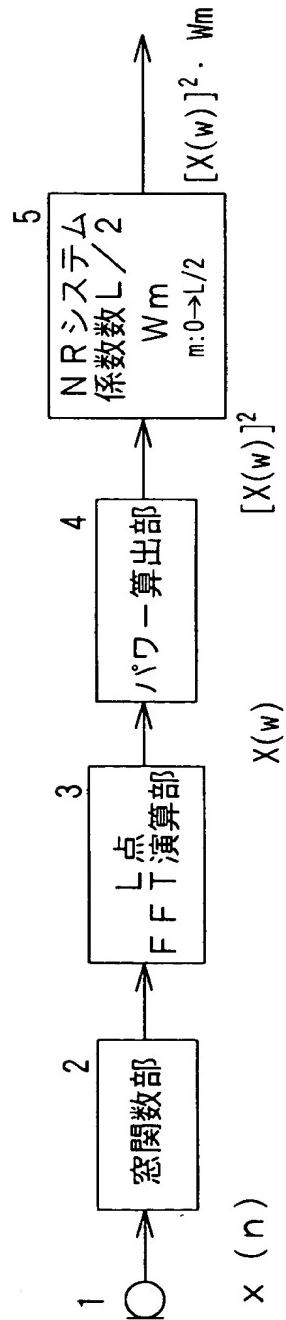
【符号の説明】

- 1 マイク
- 2 窓関数部
- 3 F F T演算部
- 4 パワー算出部
- 5 ノイズリダクションシステム
- 1 1 模擬音声発声部
- 1 2 ゲイン調整部
- 1 5 適応フィルタ
- 1 7 F F T演算部
- 1 8 パワー算出部

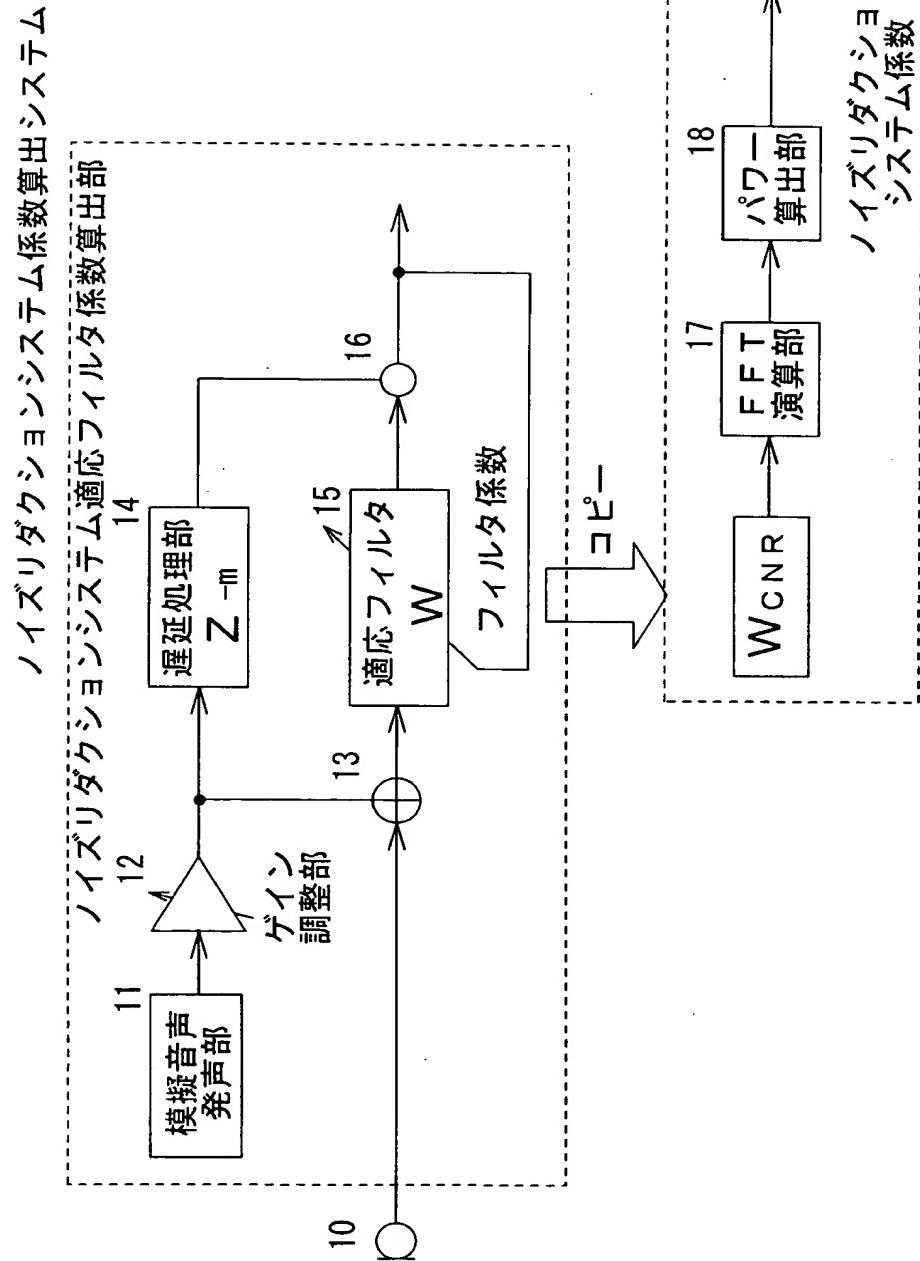
【書類名】

図面

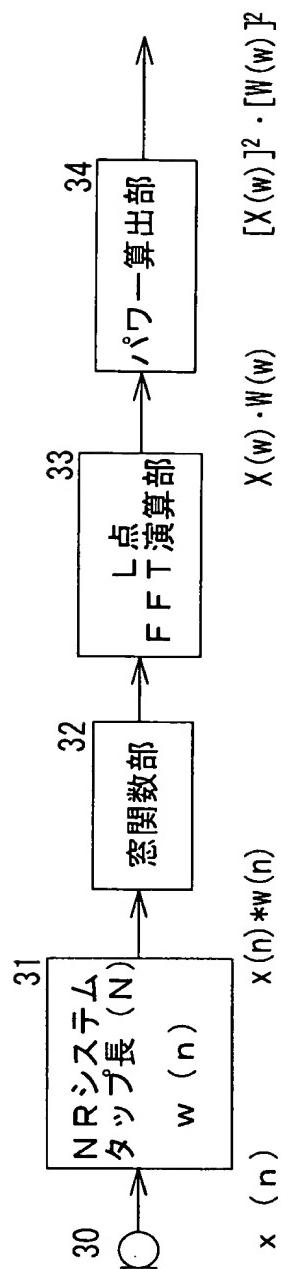
【図1】



【図2】



【図3】



【書類名】 要約書

【要約】

【課題】 音声認識装置等の、ノイズに影響されず正確に音声特徴量を抽出するための装置においては、入力音声信号に対して最初タップ長Nのノイズリダクションシステムで処理し、これをL点についてFFT処理してパワースペクトル計算を行うと、 $N \times L$ 回の乗算と、 $(N - 1) \times L$ 回の和算を行う必要がある。

【解決手段】 ノイズを含んだ音声をマイク1から入力し、窓関数処理部2で処理した後FFT処理部3でL点についてFFT処理を行う。これをパワー算出部4において入力した音声のパワースペクトルを演算する。一方、ノイズリダクションシステム5においては、予めこのシステムのフィルタ係数を求め、それを処理してノイズリダクション係数を計算しておき、このノイズリダクションシステムにより前記パワースペクトルを処理することによって、従来の装置の $1/(4N - 2)$ の処理量となり、演算装置の処理負担が減少し、処理も速くなる。

【選択図】 図1

特2000-212067

認定・付加情報

特許出願の番号	特願2000-212067
受付番号	50000881390
書類名	特許願
担当官	第一担当上席 0090
作成日	平成12年 7月13日

<認定情報・付加情報>

【提出日】 平成12年 7月12日

次頁無

出願人履歴情報

識別番号 [000101732]

1. 変更年月日 1990年 8月27日

[変更理由] 新規登録

住 所 東京都品川区西五反田1丁目1番8号
氏 名 アルパイン株式会社